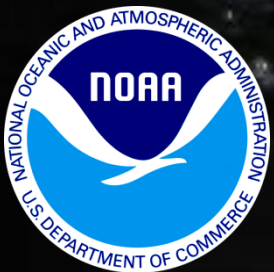




# Extending GIS with Statistical Models to Predict Marine Species Distributions



Zach Hecht-Leavitt  
NY Department of State  
Division of Coastal Resources

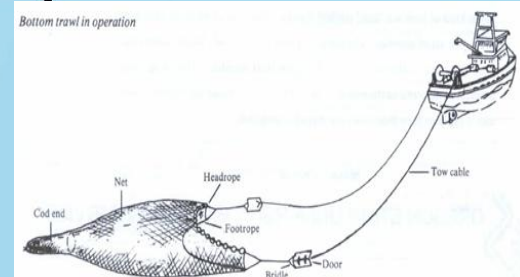


# Offshore Planning

NY



Bottom trawl in operation

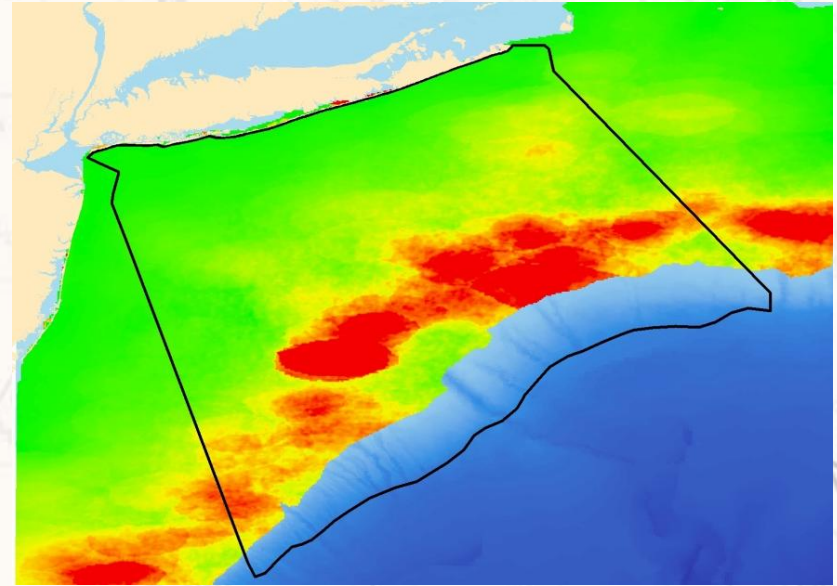
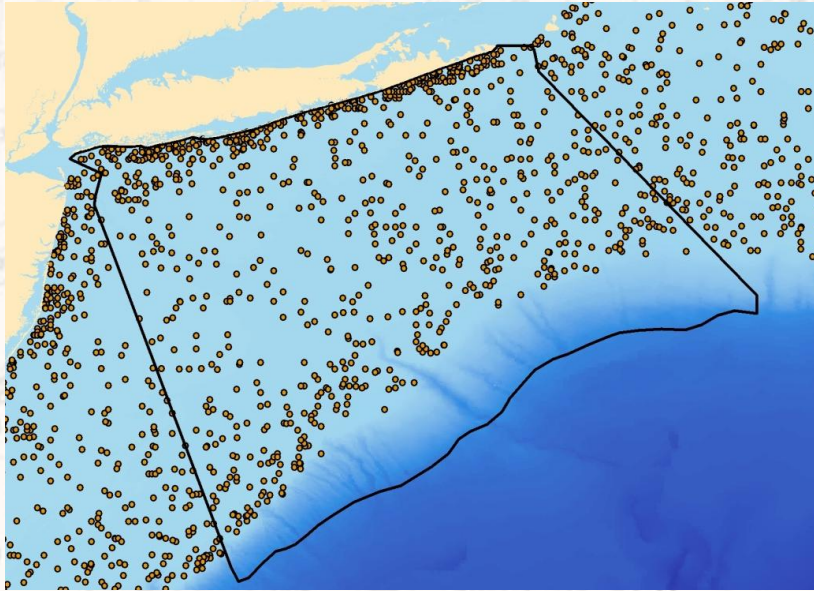




# Goals

- **Predict the abundance** of selected groundfish species as a function of:
  - **environmental/habitat variables**
  - **spatial autocorrelation**
- **Assess error**
- **Understand ecology**

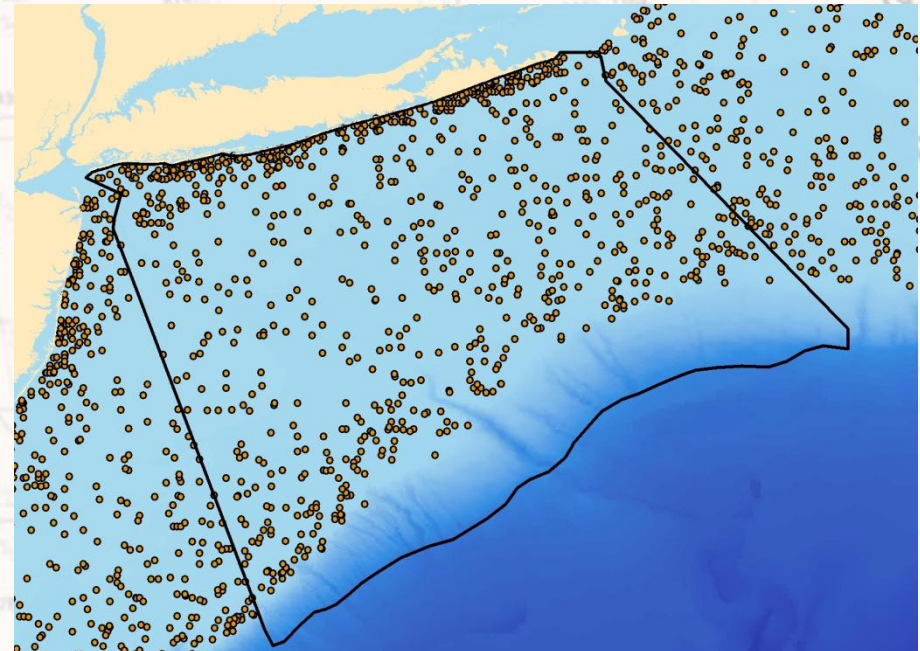
# Goals





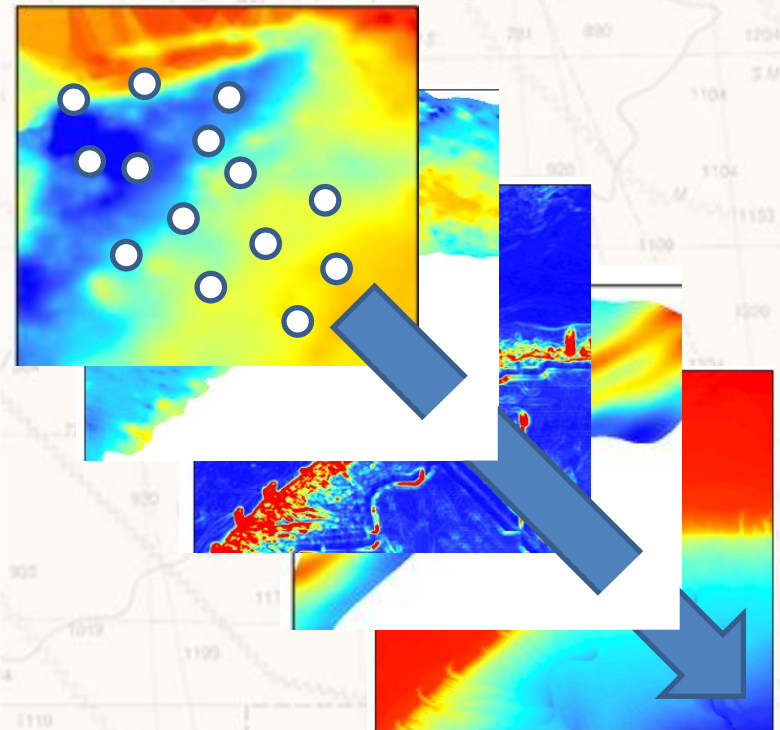
# The Fish Data

- NOAA Northeast Fisheries Science Center bottom trawl (catches **groundfish**)
- Biannual 1975-2009
- **Standardized** gear, speed, and distance
- Cleaned by Stone Environmental
- Break down by **season** and **life stage**
- 6 species selected



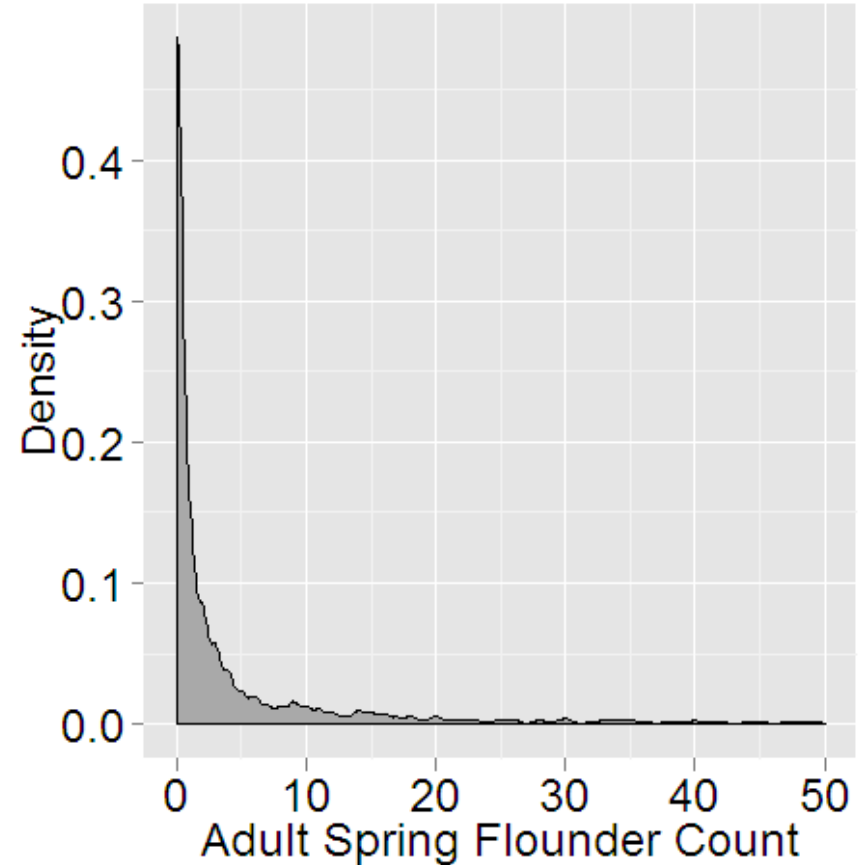
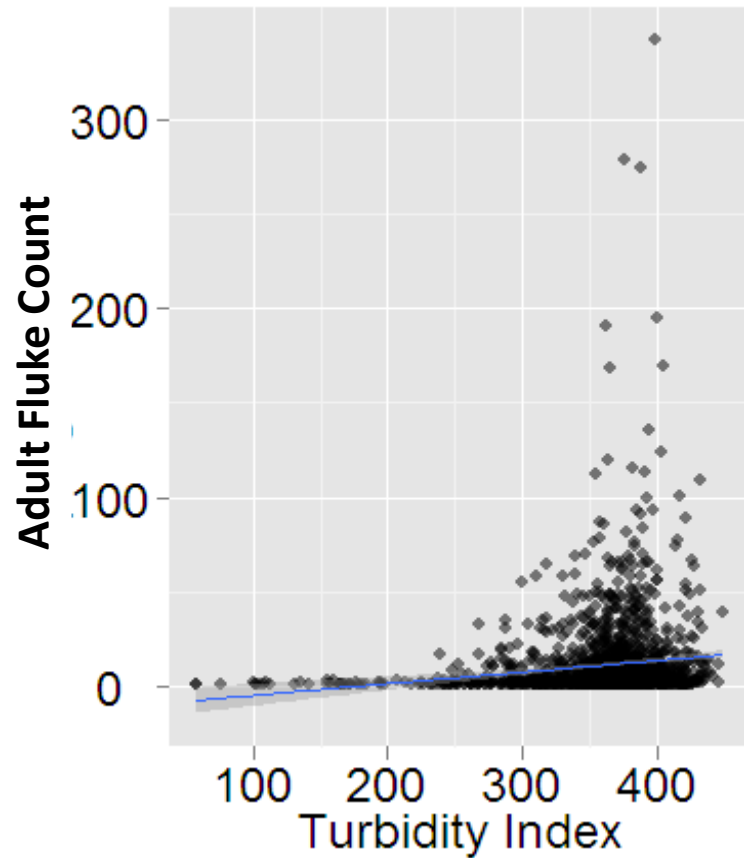
# The Environmental Data

- Depth
- Distance from Shelf Edge
- Bottom Grain Size
- Slope
- Sea Surface Temperature\*
- Chlorophyll\*
- Stratification\*
- Turbidity\*
- Zooplankton\*
- Provided by NOAA  
Biogeography Branch



\*Long-term, seasonal average

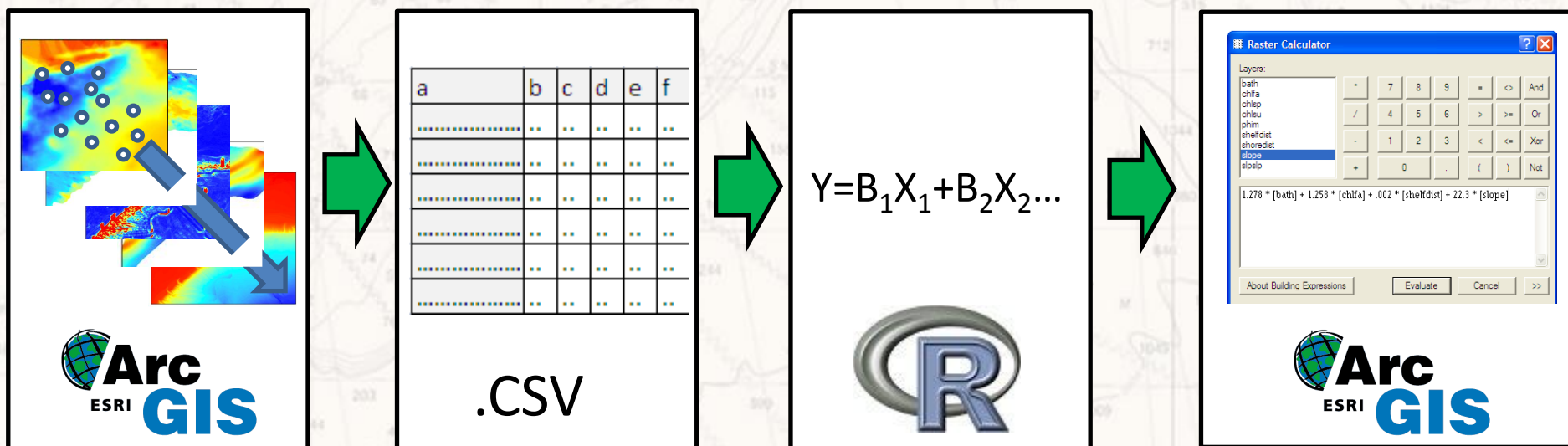
# Data Exploration



- Approximate environmental relationships with **linear trend**
- Data is **extremely skewed** (lots of zeroes)
- Go with **Zero-Inflated Generalized Linear Models (GLMs)**



# Workflow



zeroinfl(),  
Achim Zeileis

- “Loose” coupling
- With 10.x can “hard” couple via Geospatial Modeling Environment (GME)



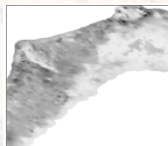
# Workflow

**B1 \***



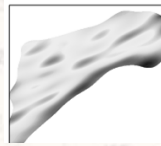
+

**B2 \***

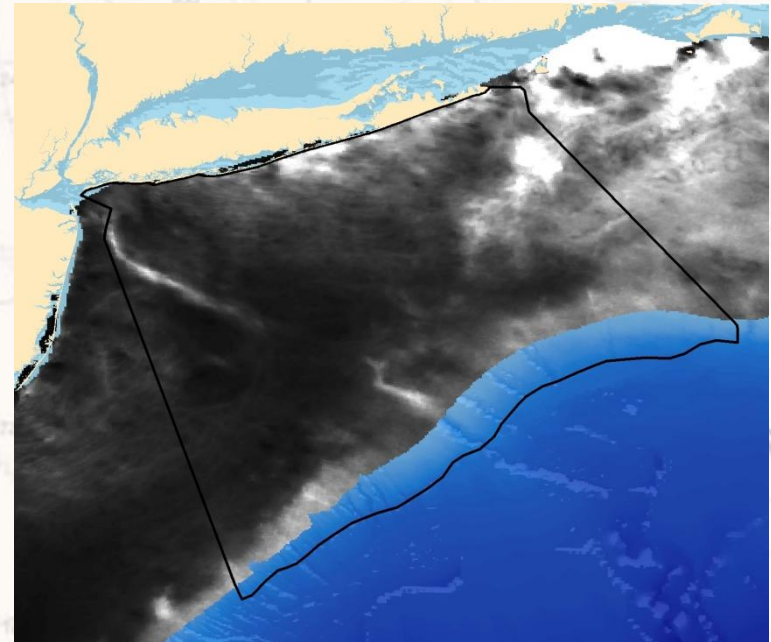


+

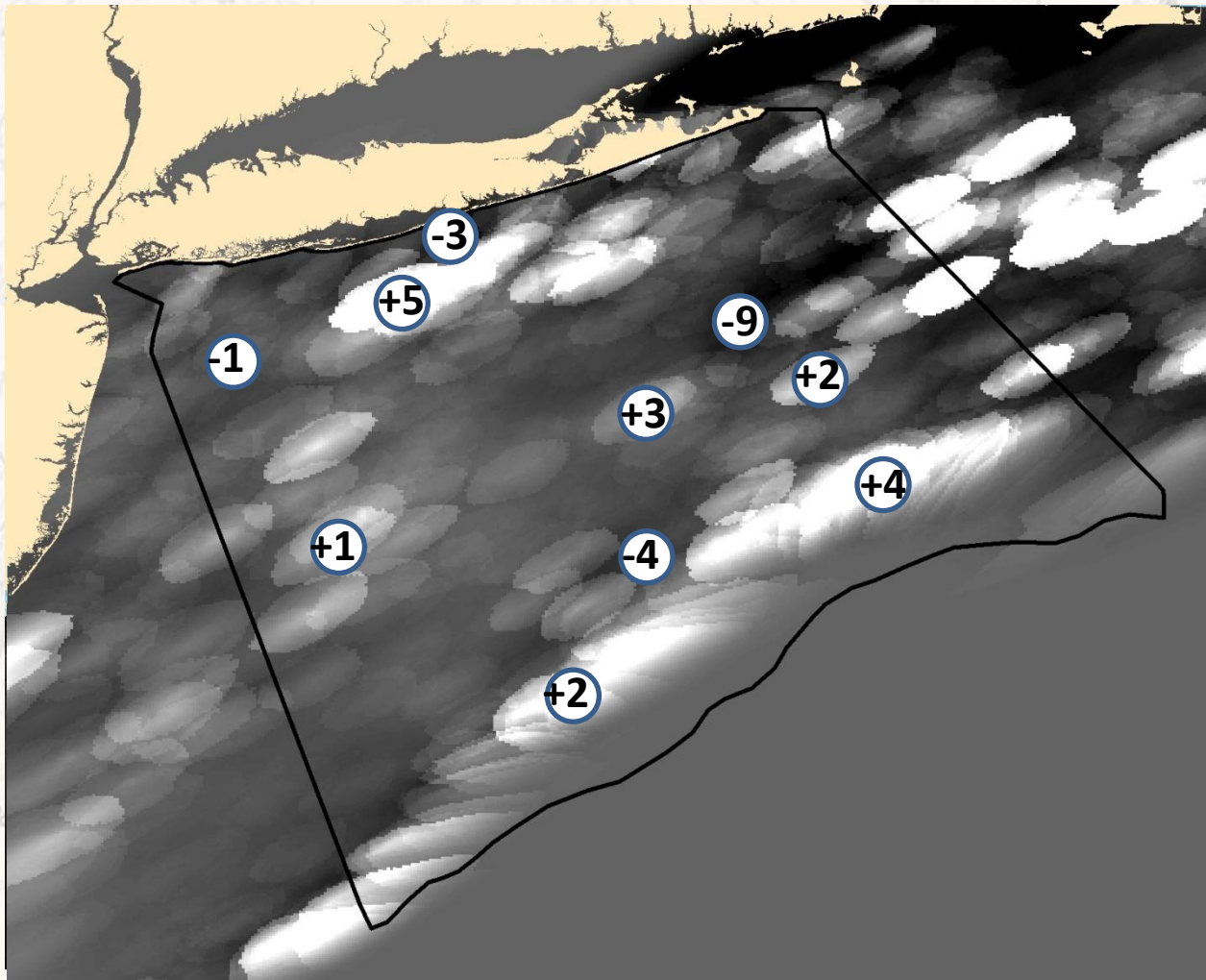
**B3 \***



=

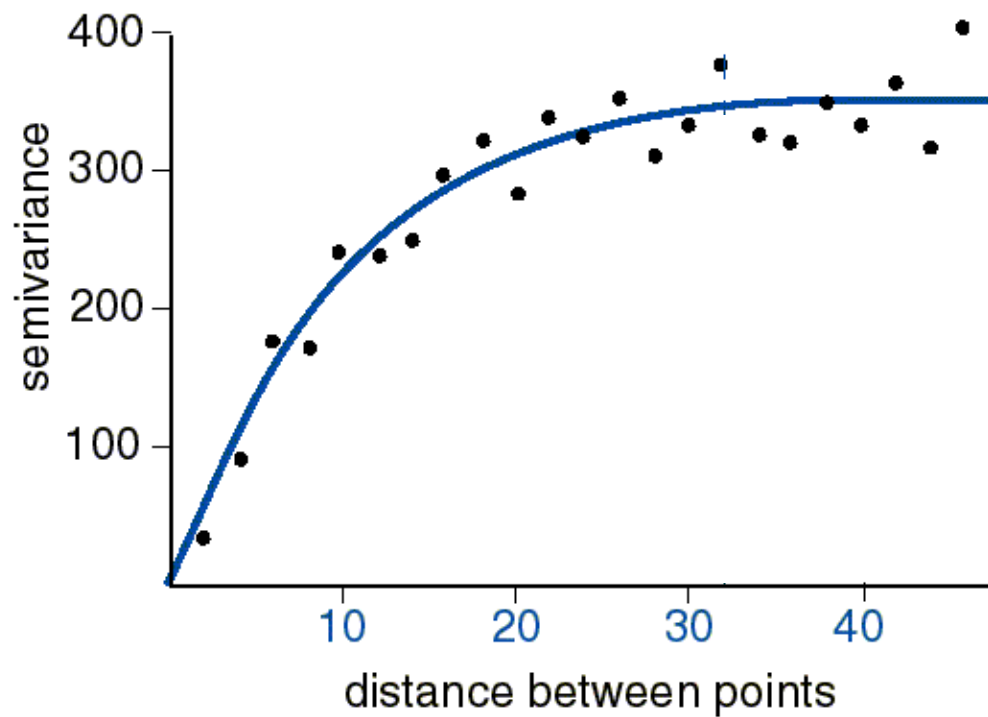


# The Residual

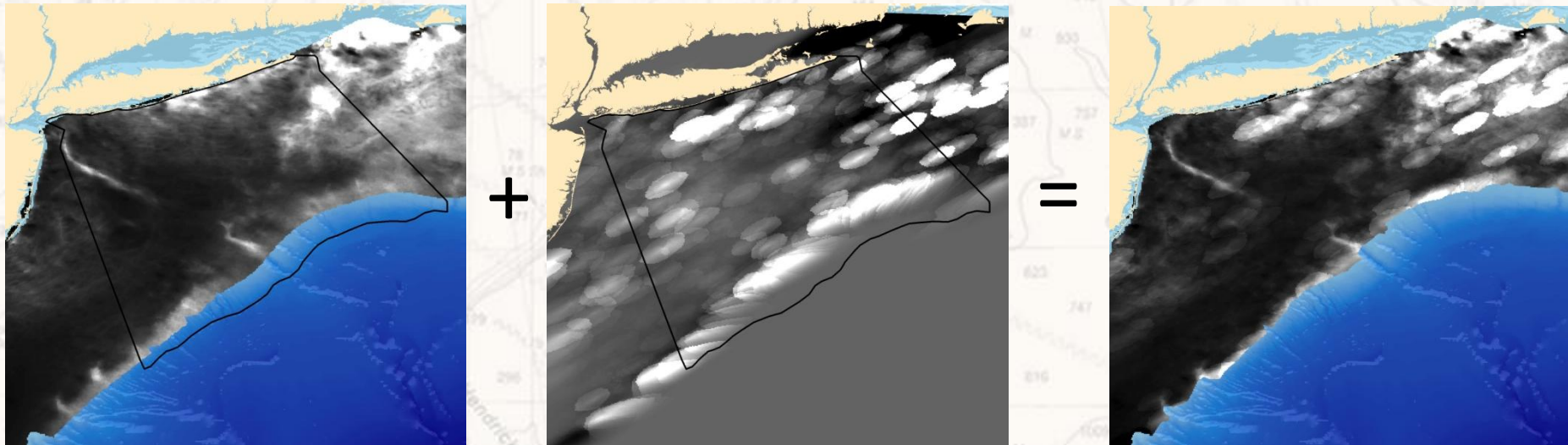




# Kriging Interpolation



# Modeling Steps



**Predictors (GLM)**

**+**

**Residual (kriging)**

**=**

**Mean Expectation**

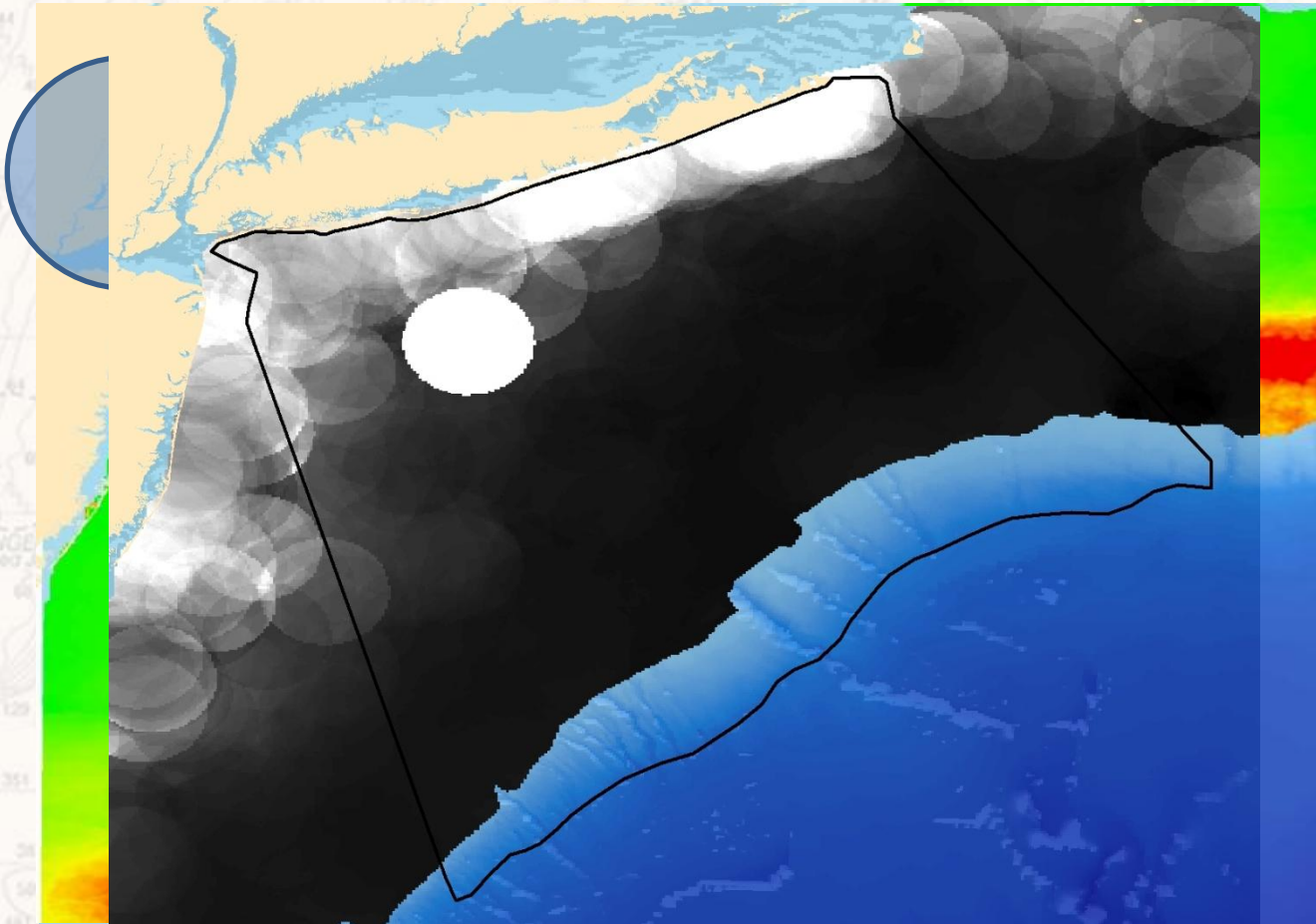


**+ Error**

**(unexplained variation)**



# Error Assessment



# Error Assessment

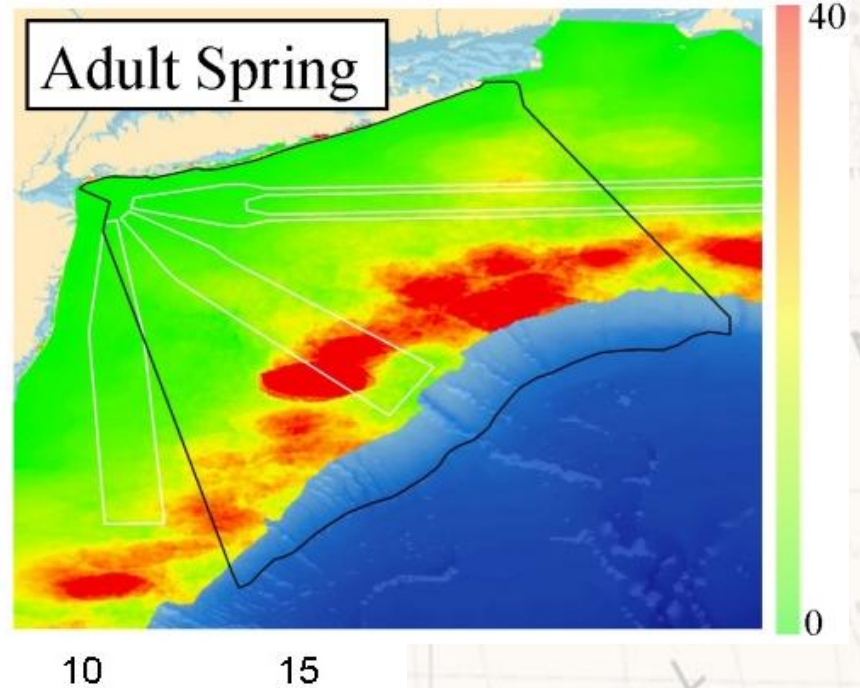
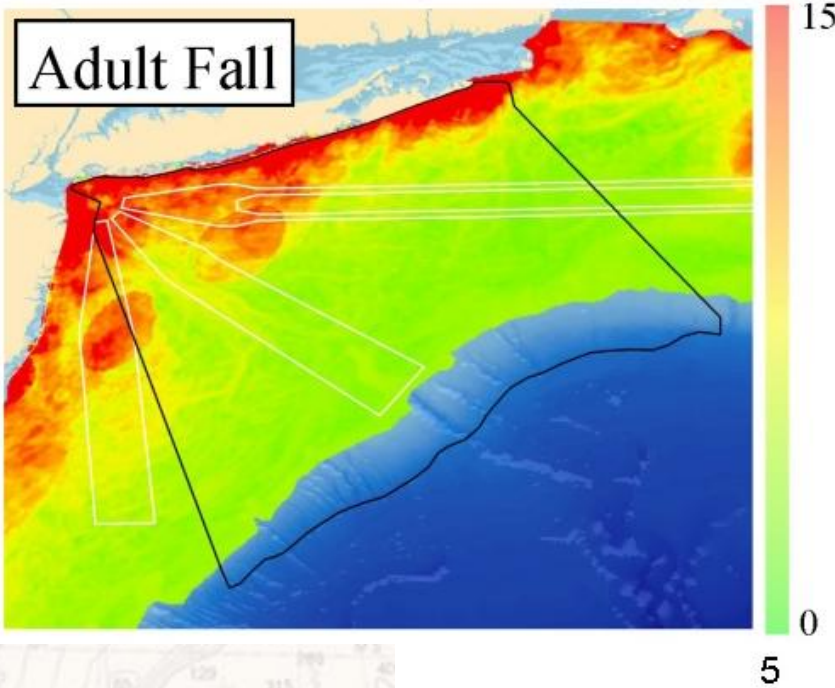
- **50/50 cross-validation**
- Smooth error with **moving window**
- **Final maps** based on **full dataset**
- **Conservative** error estimates



# Final product






(fluke example)

Adult Fall

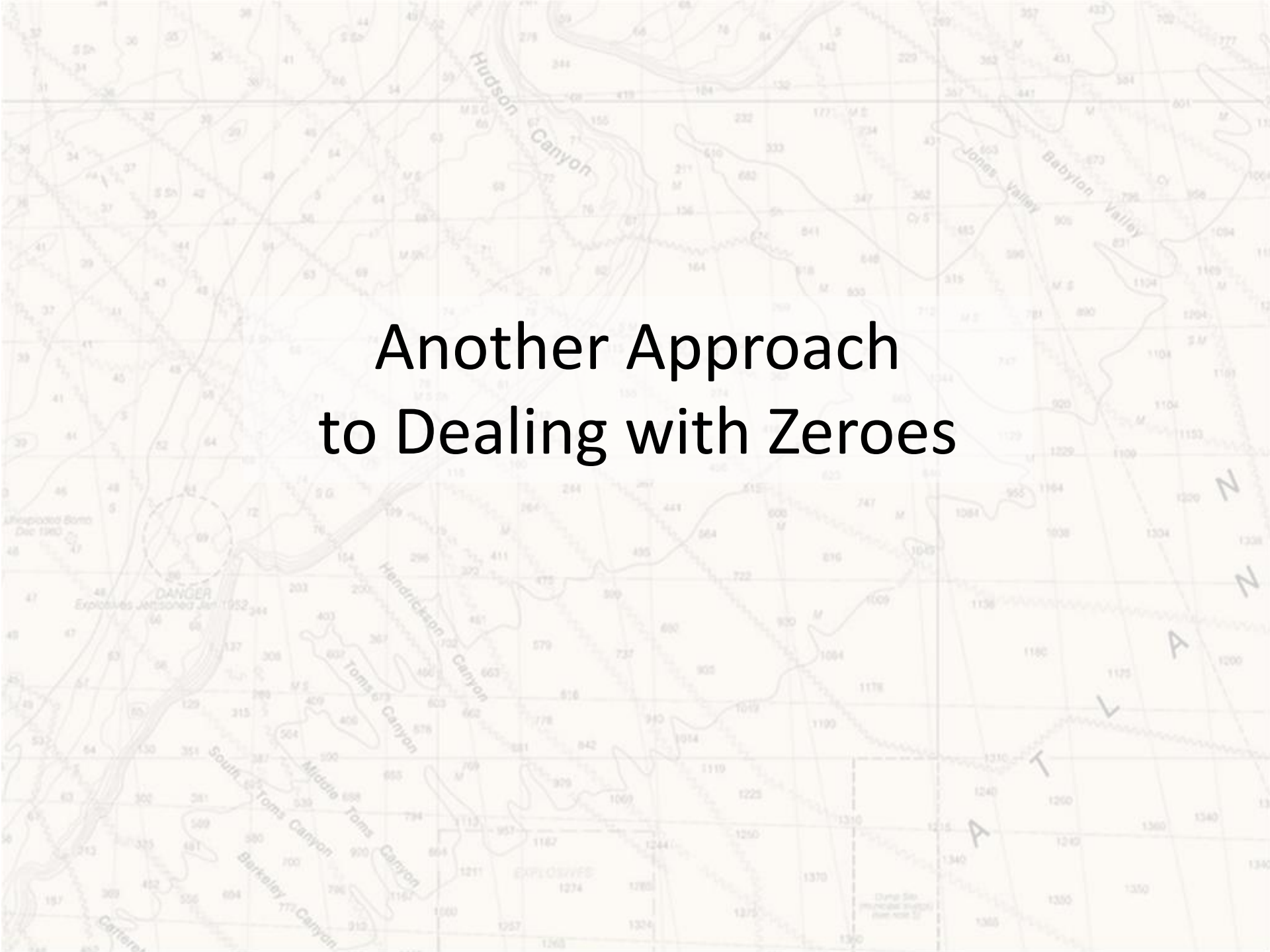


Predicted Abundance

# Goals

- **Predict the abundance of selected groundfish as a function of:**
  - **environmental/habitat variables** 
  - **spatial autocorrelation** 
- **Assess error** 
- **Understand ecology**  

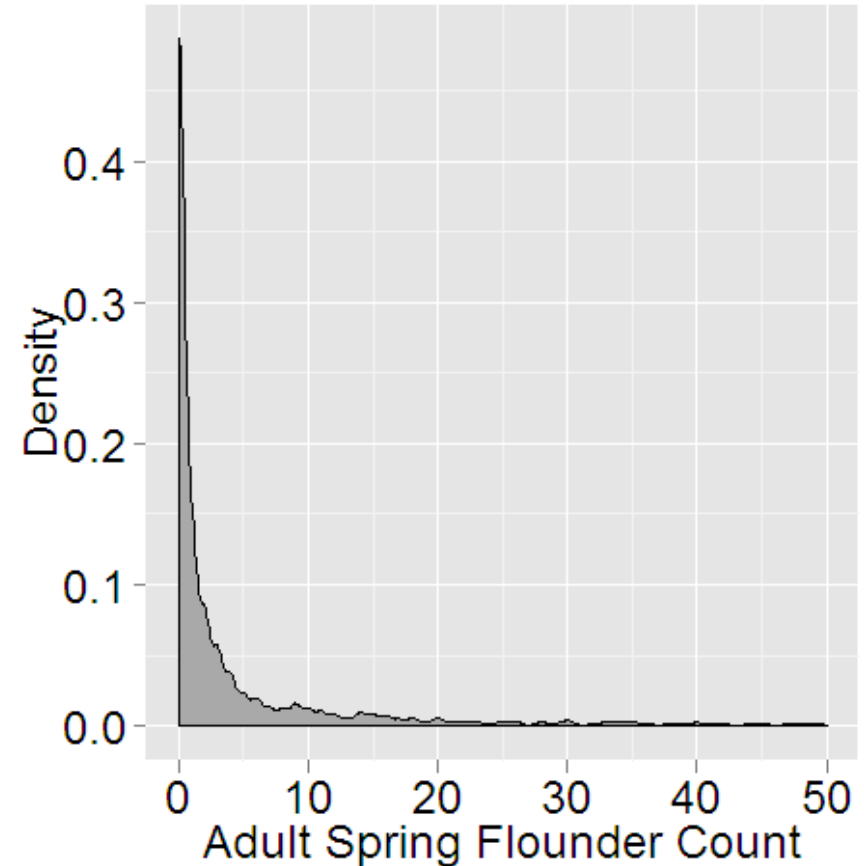


A topographic map of the Toms Canyon area, showing various canyons and valleys. A red circle highlights a 'DANGER' label near the 'Explosives Jeopardized Jan 1952' site. The map includes contour lines, elevation markers, and labels for 'Hudson Canyon', 'Jones Valley', 'Babylon Valley', 'Hendrickson Canyon', 'Toms Canyon', 'Middle Toms Canyon', 'Berkeley Canyon', and 'Carteret'. The text 'Another Approach to Dealing with Zeroes' is overlaid in the center.

# Another Approach to Dealing with Zeroes

# Another Approach

- Rather than a one-size-fits all model...
- ... model **presence/absence** and **abundance** with **two separate stages**
- May **better reflect ecological reality**
- More **conservative** approach



# Goals

- **Predict the abundance of selected cetaceans as a function of:**
  - **spatial autocorrelation**
- **Assess error**

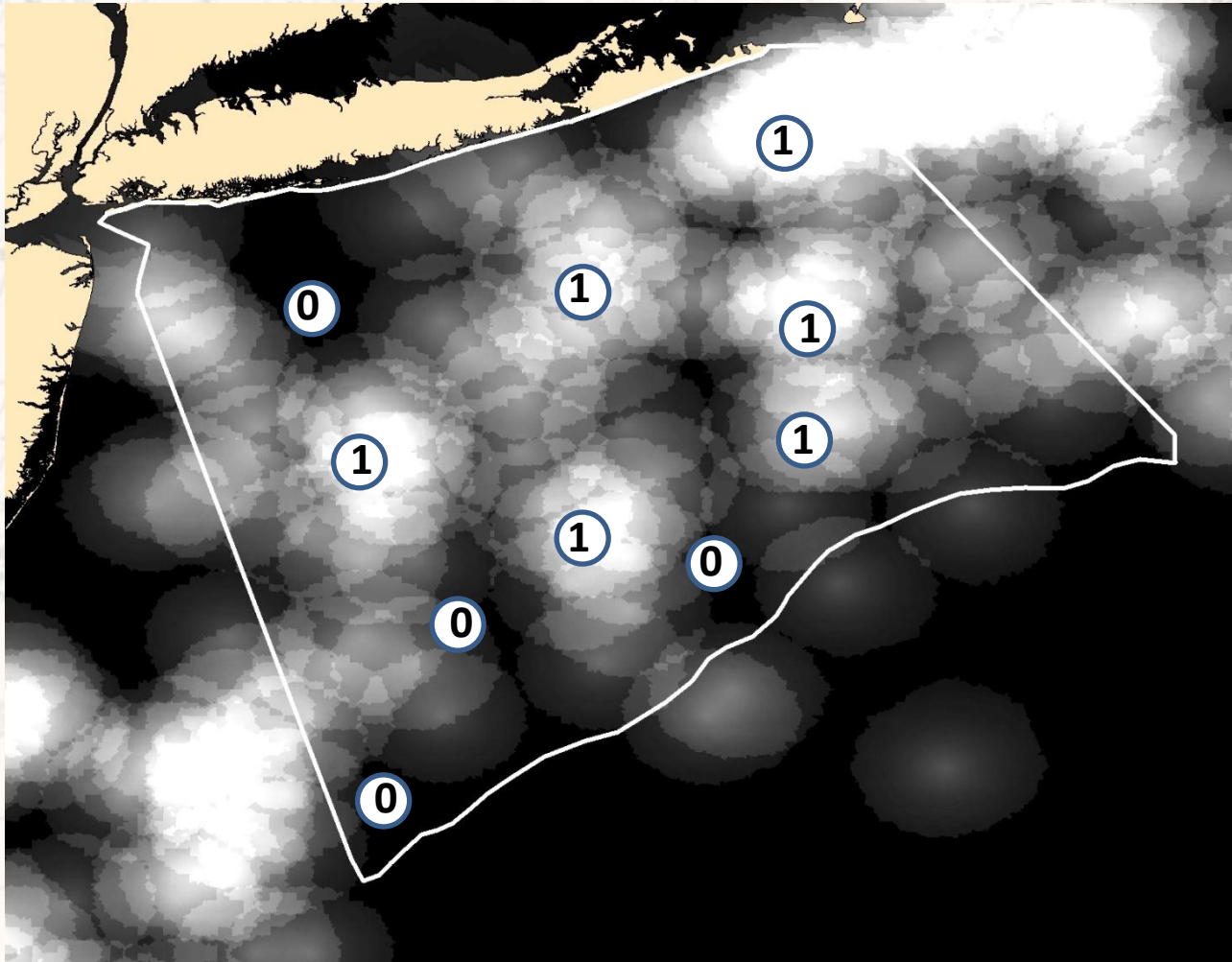


# The Marine Mammal Data

- North Atlantic Right Whale Consortium Database
- Aerial and shipboard surveys, 1978 - 2009
- Cleaned by New England Aquarium
- **SPUE**
- 4 species/groups selected
- No predictors this time!

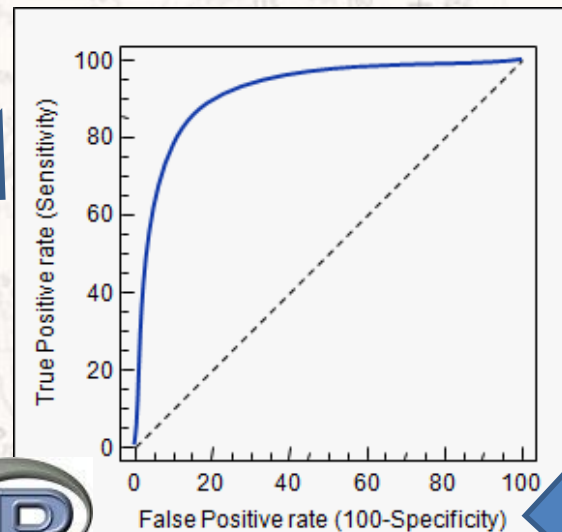
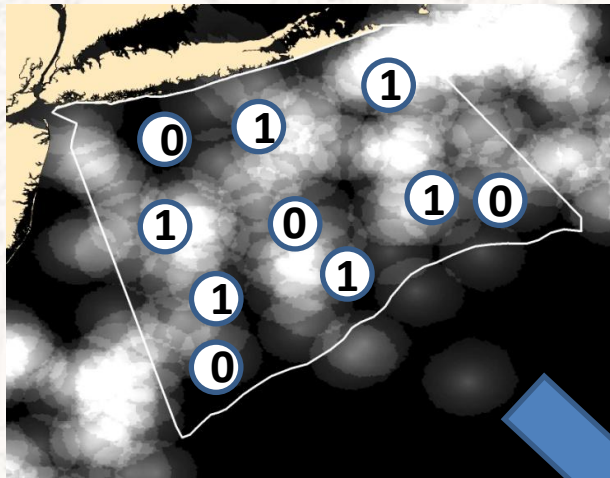


# Stage I – Presence/Absence





# Stage I – Presence/Absence



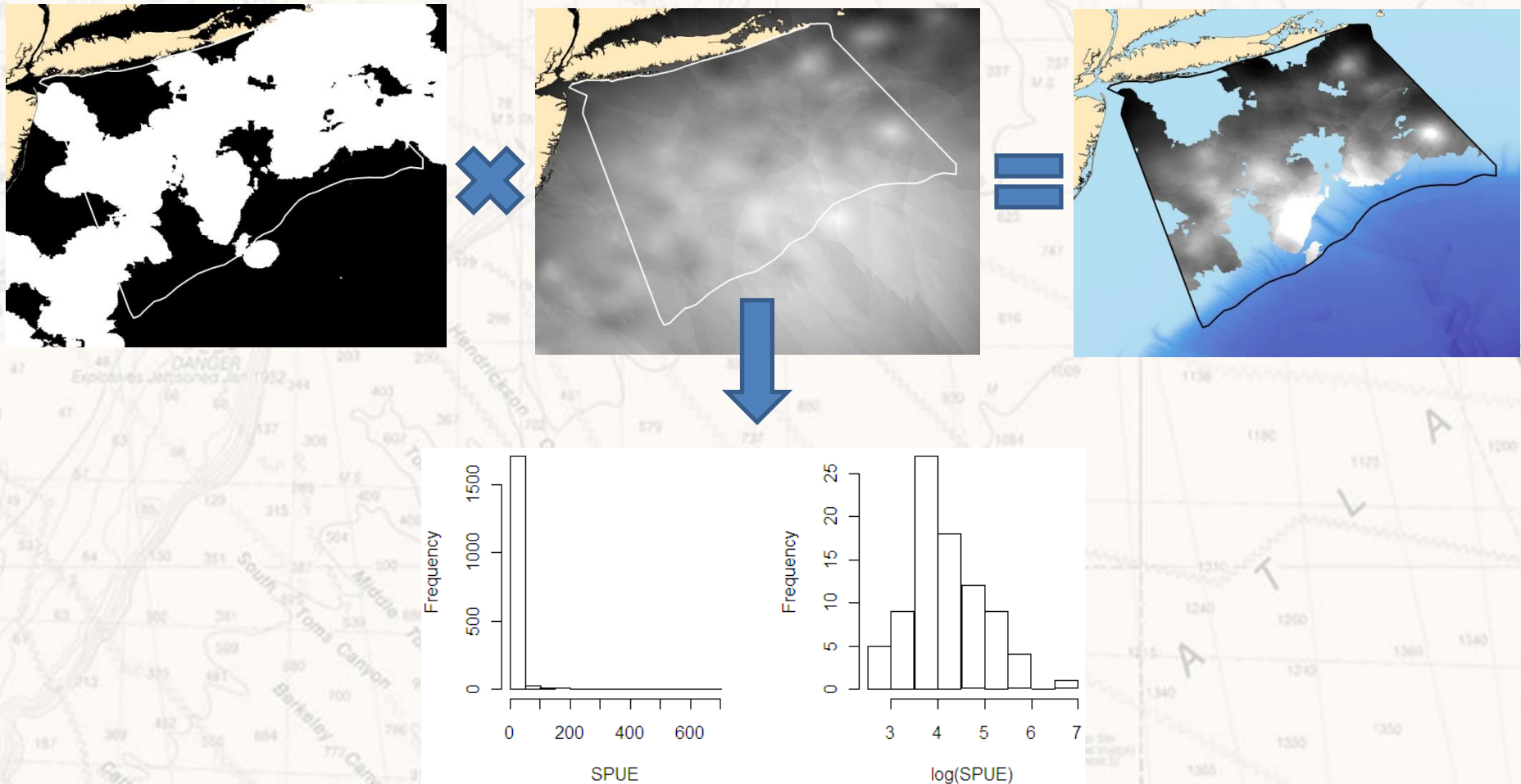
ROC Curve

Epi package,  
Carstensen et. al.

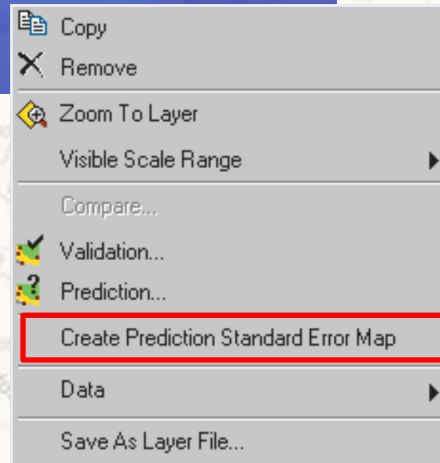
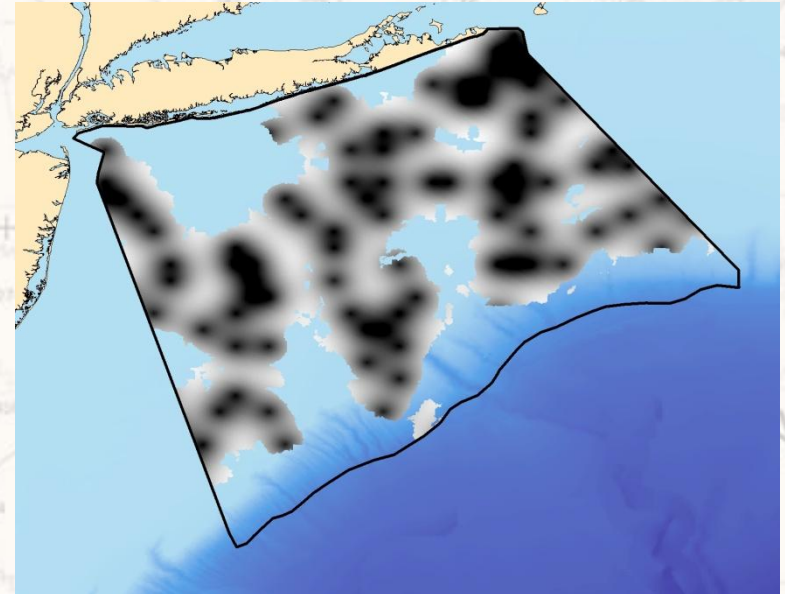
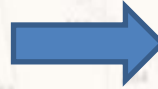
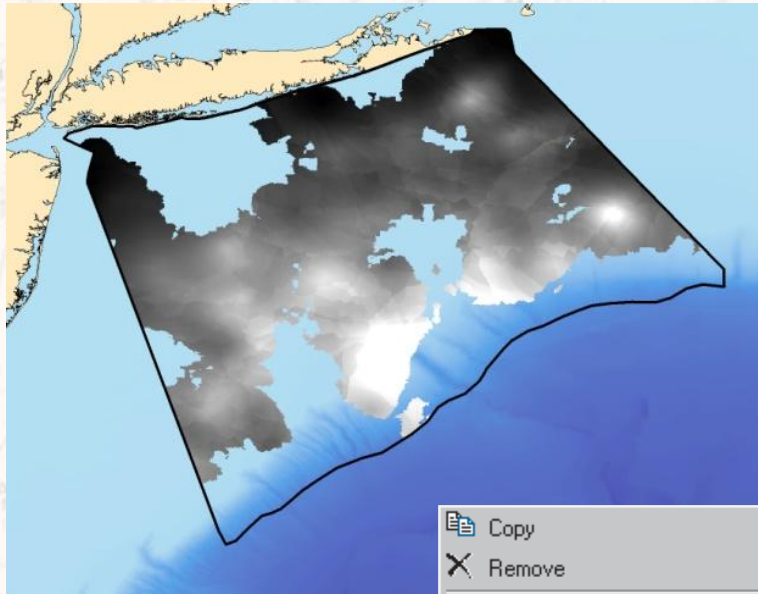






# Stage I \* II – Abundance Where Present



# Error Assessment

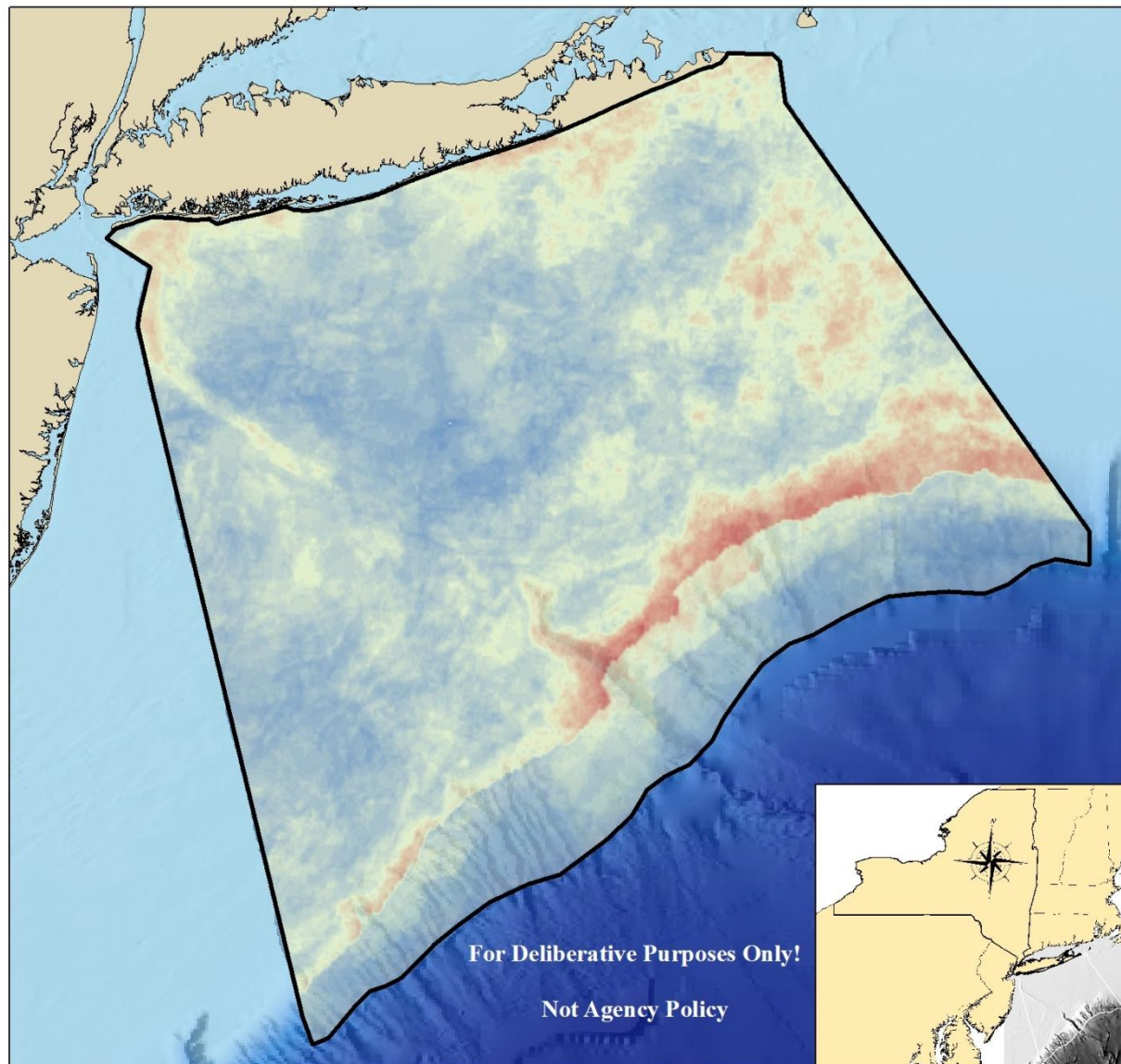


# Goals

- **Predict the abundance of selected cetaceans as a function of:**
  - **spatial autocorrelation** 
- **Assess error** 



# End goal



# Further reading

- Hengl, T., G.B.M. Heuvelink and D.G. Rossiter. 2007. **About regression-kriging: From equations to case studies.** Computers and Geosciences 33:1301-1315.
- Wenger, S. J. and M. C. Freeman. 2008. **Estimating Species Occurrence, Abundance, And Detection Probability Using Zero-Inflated Distributions.** Ecology 89:2953-2959
- Monestiez, P., L. Dubroca, E. Bonnin, J.-P. Durbec, C. Guinet. 2006. **Geostatistical modeling of spatial distribution of *Balaenoptera physalus* in the Northwestern Mediterranean Sea from sparse count data and heterogeneous observation efforts.** Ecological Modeling, 193: 615-628
- Menza, C., B.P. Kinlan, D.S. Dorfman, M. Poti and C. Caldow (eds.). 2012. **A Biogeographic Assessment of Seabirds, Deep Sea Corals and Ocean Habitats of the New York Bight: Science to Support Offshore Spatial Planning.** NOAA Technical Memorandum NOS NCCOS 141. Silver Spring, MD. 224 pp.



A topographic map of the Toms Canyon area, showing various canyons and valleys. The map includes contour lines, elevation markers, and labels for geographical features. A large, bold, black text overlay reads "Thank you!".

Thank you!

Zach.Hecht-Leavitt@dos.ny.gov / zhechtle@gmail.com